

<b>REPORT DOCUMENTATION PAGE</b>			Form Approved OMB NO. 0704-0188	
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>				
1. REPORT DATE (DD-MM-YYYY) 08-12-2014		2. REPORT TYPE Final Report		3. DATES COVERED (From - To) 28-Aug-2012 - 27-Aug-2014
4. TITLE AND SUBTITLE Final Report: Collection and Analysis of Crowd Data with Aerial, Rooftop, and Ground Views			5a. CONTRACT NUMBER W911NF-12-1-0446	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER 611103	
6. AUTHORS Mubarak Shah, Aidean Sharghi			5d. PROJECT NUMBER	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES University of Central Florida 12201 Research Parkway, Suite 501  Orlando, FL 32826 -3246			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSOR/MONITOR'S ACRONYM(S) ARO	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 61572-CS-RIP.1	
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited				
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.				
14. ABSTRACT In this report we cover purchased equipment, with their manufacturers and cost, as well as a brief description of the projects on which they were used. We have purchased aerial platforms, imaging systems, and flight control systems and used them in different projects that will be discussed in the following.				
15. SUBJECT TERMS				
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	15. NUMBER OF PAGES
a. REPORT UU	b. ABSTRACT UU	c. THIS PAGE UU		
			19a. NAME OF RESPONSIBLE PERSON Mubarak Shah	
			19b. TELEPHONE NUMBER 407-823-5077	



## Report Title

Final Report: Collection and Analysis of Crowd Data with Aerial, Rooftop, and Ground Views

### ABSTRACT

In this report we cover purchased equipment, with their manufacturers and cost, as well as a brief description of the projects on which they were used. We have purchased aerial platforms, imaging systems, and flight control systems and used them different projects that will be discussed in the following.

---

**Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:**

**(a) Papers published in peer-reviewed journals (N/A for none)**

Received

Paper

**TOTAL:**

Number of Papers published in peer-reviewed journals:

---

**(b) Papers published in non-peer-reviewed journals (N/A for none)**

Received

Paper

**TOTAL:**

Number of Papers published in non peer-reviewed journals:

---

**(c) Presentations**

Number of Presentations: 0.00

---

**Non Peer-Reviewed Conference Proceeding publications (other than abstracts):**

Received      Paper

**TOTAL:**

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

---

**Peer-Reviewed Conference Proceeding publications (other than abstracts):**

Received      Paper

**TOTAL:**

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

---

**(d) Manuscripts**

Received      Paper

**TOTAL:**

Number of Manuscripts:

Books

Received      Book

TOTAL:

Received      Book Chapter

TOTAL:

Patents Submitted

Patents Awarded

Awards

Graduate Students

<u>NAME</u>	<u>PERCENT_SUPPORTED</u>
FTE Equivalent:	
Total Number:	

Names of Post Doctorates

<u>NAME</u>	<u>PERCENT_SUPPORTED</u>
FTE Equivalent:	
Total Number:	

---

### **Names of Faculty Supported**

NAME

PERCENT SUPPORTED

**FTE Equivalent:**

**Total Number:**

---

### **Names of Under Graduate students supported**

NAME

PERCENT SUPPORTED

**FTE Equivalent:**

**Total Number:**

---

### **Student Metrics**

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: ..... 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:..... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:..... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):..... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:..... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense ..... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields: ..... 0.00

---

### **Names of Personnel receiving masters degrees**

NAME

**Total Number:**

---

### **Names of personnel receiving PHDs**

NAME

**Total Number:**

---

### **Names of other research staff**

NAME

PERCENT SUPPORTED

**FTE Equivalent:**

**Total Number:**

---

**Sub Contractors (DD882)**

**Inventions (DD882)**

**Scientific Progress**

\*See attachment\*

**Technology Transfer**

# **DURIP: Collection and Analysis of Crowd Data with Aerial, Rooftop, and Ground Views**

Final Report  
11/10/2014

Mubarak Shah  
Center for Research in Computer Vision (CRCV)  
University of Central Florida



This report covers equipment purchased under this project, along with information about manufacturers and cost, as well as a brief description of the projects in which they were used. We have purchased aerial platforms, imaging systems, and flight control systems and used them in different projects described next.

## **1. Dataset Collection**

### **1.1. Dataset collection using UAS**

Datasets are important for evaluation of scientific methods and techniques. As a part of our ongoing research on Unmanned Aircraft Systems (UAS) and high precision Imaging Systems, we collected datasets to capture and evaluate real scenarios instead of synthetic and simplified ones. We collected these datasets using different aircrafts.

Erista 8 HL OctaCopter is a heavy-lift aerial platform capable of using high-resolution cinema-grade imaging systems. This UAS has a 3-Axis servo stabilized camera gimbal mechanism to ensure a steady shot in all environments and flying conditions. This mechanism eliminates the hassle of stabilizing captured videos, making the real-time application feasible. We equipped this aircraft with 4K BlackMagic Production Camera that has the ability to record video at a high data rate. This enabled us to collect a dataset of densely populated scenes with extreme motion and dynamic range requirements that can be used to demonstrate the feasibility and performance of tracking algorithms on videos captured through UAS.

Venu 4 QuadCopter is a 2-Axis gimbal equipped UAS. It has a built-in GPS stabilizer. Due to its compact size, it is capable of capturing First-Person-View (FPV) and close quarters flying. However, it can only carry smaller cameras compared to the OctaCopter described above. A dataset of crowd videos was captured using this UAS for the purpose of human detection and tracking.

Two different flight control systems were purchased in order to control the above UAS. Pixhawk with Telemetry Link is a sophisticated flight control system suitable for all UAS platforms. The flexibility of the Pixhawk platform enables high level of autonomous flight; in connection to the telemetry link, remote deployments with preplanned navigational data can be achieved with high

levels of precision. More features are available using additional sensors that allow the controller to track, detect and follow objects with reasonable precision.

It is worth noting here that in order to collect these datasets, our team underwent training to learn to fly these aircrafts. Safety was of key significance and a thorough preflight checklist and maintenance program was put into effect.

## **1.2 Dataset collection using 4K cameras**

AVT ProSilica GX3300C is a high-resolution ultra HD camera (4K) that was used to collect a dataset in downtown Orlando. This dataset features high-resolution traffic flow with oblique view angle, which introduces new challenges for automated video analysis. There are several unique factors associated with oblique-view video. In particular, due to perspective effects, size of vehicles and their speeds vary drastically. Also occlusions happen more often than in videos captured at other angles.

Blackmagic Production 4K is another high-resolution camera that is cinema grade and high quality, with the capability of capturing videos with 4K resolution at 30 frames per second. It can capture RAW format so we are able to perform color correction in post-production. This is helpful in computer vision in making the boundaries clearer and improves detection and tracking consequently.

An Infra-Red (IR) camera is also used in conjunction with 4K cameras to provide additional data. By adding an IR camera, we supplement data with another modality that can provide very useful information after registration. This additional data can be used to provide depth information and also improve quality of both modalities in their shared patches of image.

## **2. Projects**

All these efforts have two main reasons. First, we intend to show how current technology can help computer vision researchers. Second, our goal is to collect datasets close to real scenarios on which methods and techniques developed from computer vision problems can be tested.

In the following, we briefly discuss the projects in which the collected datasets were used.

## 2.1 Crowd counting using 4K cameras

The intention of this project was to determine if improved image quality improves the accuracy of existing methods. Initial experiments were performed using the crowd counting work [1] from UCF CRCV group in this brief project. We captured a video of a graduation ceremony using the BlackMagic Production 4K camera and another Canon high-definition (HD) camera simultaneously. We then ran the method on a few manually selected frames of the video. Higher dynamic range and availability of RAW format for the frames in the BlackMagic camera resulted in better accuracy, however, since the resolution of each frame is almost four times that of the Canon HD, it requires more time to process it as well. 4K camera easily outperformed the HD camera in the dense areas and darker patches. In the rest of patches, frames from both cameras give decent results. In the following we show some pictures obtained and output of the applied method.



*Results from BlackMagic Camera. Count=752. Error=10.58%*



*Results from Canon 1080p Camera. Count=520. Error=23.53%*

To show better where the error initiates in regular 1080p camera, we zoomed in so we can clearly see why 4K camera is reducing the error by 10 percent.



*The picture on the left is cropped from the images shown above. Left: BlackMagic. Right: Canon 1080p.*



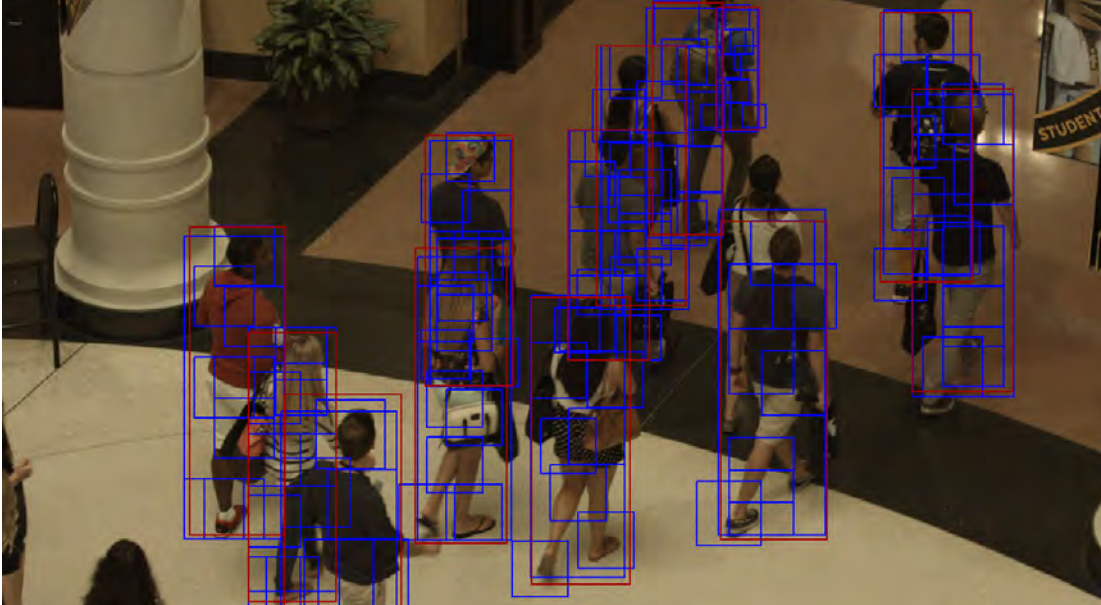
*The above images show the heads marked by the method on both images.*

As we can see, due to higher details available in 4K BlackMagic camera, the method is giving us a more accurate count of heads.

## **2.2 Human Detection and Tracking using 4K cameras**

For the tasks of human detection and tracking, similar to the previous project, we captured sequences of video with two cameras, one 4K BlackMagic camera and an HD camera. These sequences were gathered in the Student Union at UCF. In these clips students are walking unaware of the camera so the setup is a real scenario. We applied DPM [2] human detection method on clips, and the 4K camera outperformed the HD camera in this scenario as well, with a large margin in detection and subsequently in tracking. Detections were more robust and ID switch rate was much lower compared to the HD sequences. To identify the reasons for these differences, we have studied the misdetections and false positive/negatives in HD sequences

along with 4K versions of them. First, Deformable Part Model (DPM) method was able to better locate parts of the human body due to higher resolution and second, the higher dynamic range resulted in more crisp boundaries which enables the DPM method to distinguish people walking together. Due to better detection, tracking which associates detections in consecutive frames is much more successful in 4K videos.



*This image shows output of DPM [2] on a frame captured using BlackMagic 4K camera.*

### 2.3 Feasibility study of Search and Rescue using UAS

The goal in this project was to study the effectiveness of a UAS in emergency situations. So a mock scenario was setup to identify a lost citizen in a remote location. To do so, the UAS was fit with wireless video downlink, the video was used to find the person in distress and pinpoint the location. This location identification is useful for emergency services to provide assistance and improve response time, especially due to the fact that no human supervision is required. This project is still ongoing at CRCV UCF and uses UAS, imaging system, and a flight control system. The battery type primarily determines the effective flight range of the UAS system. Typical runtime with high capacity batteries is roughly 20 minutes. Also, Zeiss Cinemizer OLED was used to provide a suitable display mechanism to view real time video transmitted from the UAS. The following figure shows two frames from the sequence captured for this project.



## **2.4 Real-Time Crowd Behavioral Analysis and Crime Detection using Network Surveillance cameras**

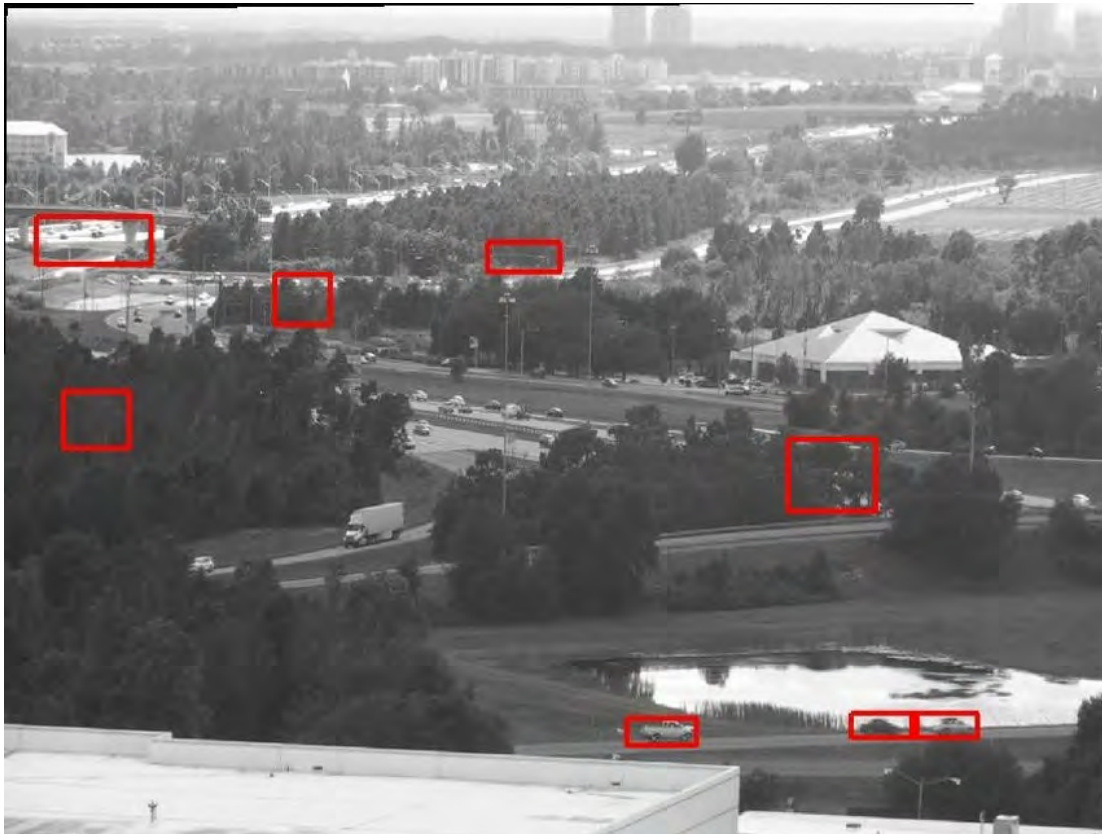
Axis 5534-E Network Camera provides surveillance data in extremely crowded scenes. It also has large optical zoom that helps capture high levels of detail when zoomed in. The goal in this project was to analyze crowd behavior, for instance, segmenting the motions in the scene, and use that to study the crowd behavior and group formation. This information can be used in higher-level projects, such as crime detection in crowds.

Furthermore, we are working to find an automated way to analyze the video in real-time and identify any crime taking place in the scene. This can be done by analyzing the movement of the crowd in the scene and finding suspicious spontaneous activities done by individuals. For instance, if a person is standing still while the crowd is moving in a certain direction, this person

has to be marked for further analysis. If such a person exists in the scene and starts to move in the opposite direction of the crowd or moves spontaneously, it should be flagged for inspection.

## 2.5 Oblique-View Detection and Tracking

In this project, we used a dataset collected in downtown Orlando to define a new and challenging problem. Detection and tracking methods work under either in UAV circumstances or first person views. However, a much more difficult task is to detect and track objects in oblique views that add complexities for visual analysis. For this, we applied two different detection methods to observe where they fail. In the following, we show pictures of applied methods on frames obtained from the dataset we collected.



The above figure is obtained by applying Deformable Parts Model [2] trained for vehicles. Notice, except for three cars relatively close to camera, no other car is detected, and there are many false positives and negatives. The reason is that the vehicle sizes vary drastically, and there is no optimal parameter to tune the method for frames with oblique view.

We also tried a motion based detection method that is from the CRCV group and the result is the figure shown below. This method fails because there are groups of cars moving together which results in frequent merged detections.



To obtain better and more robust detections, we started using KLT tracks and develop a prototype method which we are still working on. The results of applying the prototype are shown in the following figures:

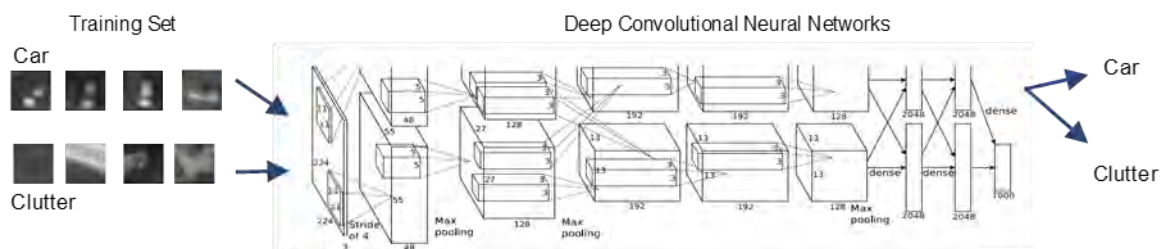




## 2.6 Aerial Vehicle Detection through Deep Convolutional Neural Networks

Aerial vehicle detection has remained an open area of research for many years. The reason is difficulty in obtaining a model representation of the objects given the small images of the vehicles. Also, obtaining features that could be discriminative enough from such images is very challenging.

In this project we have explored the use of Deep Convolutional Neural Networks (DCNN) to train a network with a set of samples of aerial views of cars. We use a modified version of the AlexNet network that adapts to our dataset and size of images better. The following figure shows our framework.



Our training and validation datasets are composed of patches of images extracted from the LAIR dataset annotated as car and clutter, as seen in the above figure.

We have also fine-tuned parameters, such as learning rates, to obtain the best possible learned weights for the model. Finally, we tested our model with images from the LAIR dataset to validate our results.

Our results indicate a significant improvement over the state of the art, and, in some cases, the results of our trained network are almost perfect. Our method indicates an average accuracy of 95.01% on our validation dataset that consists of patches obtained from three tiles from the LAIR dataset.



*An image from the LAIR Dataset.*

### III. List of Purchased Items, Costs and their Manufacturers

A detailed list of all the equipment purchased together with their costs, and the names of the vendor follows, along with their functionality in the above projects:

Product Description	Project	Functionality	Vendor	Amount
<b>Aerial Platforms</b>				
<b>Penguin B Airframe</b>				
Penguin B airframe with servos	Aerial Platform	Long flight-time aerial platform for imaging applications	UavFactory Ltd	\$15,885.00
80W Onboard Generator upgrade		Power source of running aerial computing applications	UavFactory Ltd	\$3,744.00
Pitot-Static Assembly		Pressure sensor to obtain airspeed to enable autoland and altitude hold	UavFactory Ltd	\$441.00
Header tank point level sensor			UavFactory Ltd	\$522.00
Push-Pull connector upgrade			UavFactory Ltd	\$378.00
Heavy-duty landing gear upgrade			UavFactory Ltd	\$1,152.00
Catapult hardpoints integrated		Mounting points to enable catapult launch	UavFactory Ltd	\$1,143.00
Shipping, handling, insurance			UavFactory Ltd	\$700.00
<b>Erista 8 Heavy Aerial Lift Platform</b>				
Erista 8 Standard	UAS Dataset Collection	Octacopter platform fitted with 3-Axis gimbal for high resolution aerial data acquisition	Hoverfly Technologies, Inc.	\$21,600.00
<b>HoverFly Venu Aerial Platform</b>				
SkyHero X4 Quad UAV	UAS Dataset Collection	Quadcopter platform for aerial surveillance	Hoverfly Technologies, Inc.	\$3,000.00
<b>Flight Controllers</b>				
DJI Ace One for UAV	UAS Dataset Collection	Advanced flight control system for helicopter stabilization	DJI USA	\$3,200.00
3DR Pixhawk Autopilot System	UAS Dataset Collection	Autopilot system with telemetry data link	3drobotics.com	\$582.94
3DR Pixhawk Autopilot System	UAS Dataset Collection	Autopilot system for multiple platforms	3drobotics.com	\$292.58
<b>Imaging Systems and Accessories</b>				
Blackmagic Production Camera 4	Crowd Counting using 4K Cameras	High resolution cinema grade digital video acquisition platform	Adorama	\$2,621.65
Blackmagic Production Camera 4	Human Detection and Tracking using 4K Cameras	High resolution cinema grade digital video acquisition platform	Adorama	\$2,099.75
Blackmagic Production Camera 4	Crowd Counting using 4K Cameras		Adorama	\$2,610.45
Blackmagic Production Camera 4	Human Detection and Tracking using 4K Cameras		Adorama	\$3,018.71
Blackmagic Production Camera 4			Adorama	\$7.50
Prosilica power supply and mount		Power supply for Prosilica camera system	B&H Photo Video	\$668.00

IR Camera Computer			Computer Warehouse	\$850.00
IR Camera Computer			Computer Warehouse	\$584.00
Video Card			Computer Warehouse	\$98.00
Freight			1st Vision Inc	\$44.00
Dalsa 4k LS camera 9khz		Line scan camera that enables high resolution imaging applications	1st Vision Inc	\$2,945.00
Dalsa 4k LS camera 9khz		Line scan camera that enables high resolution imaging applications	1st Vision Inc	\$2,989.00
Dalsa 2k LS camera 18khz		Line scan camera that enables high resolution imaging applications	1st Vision Inc	\$2,795.00
Camera Link Frame Grabber		Camera Link Frame Grabber for Line Scan Cameras	National Instruments	\$1,635.18
Camera Link I/O Extention Board			National Instruments	\$302.83
Camera Link Frame Grabber & Extension Board		Camera Link Frame Grabber for Line Scan Cameras	National Instruments	\$4,047.72
Camera Link I/O Extension Board			National Instruments	\$302.80
High Performance Camera Link Frame Grabber			National Instruments	\$2,140.45
Ethernet Modules 2-CH and 4-CH		Gigabit ethernet network interface to enable high speed capture from Gig.E Cameras	Mouser Electronics	\$556.28
Basler Ace, Tripod Mount, Power Adapter		Accessories for Basler Ace Camera	graftek	\$1,537.96
Axis Communications Wall Bracket	Realtime Crime Detection and Tracking	Mounting hardware for network camera	B&H Photo Video	\$96.95
Axis Communications Network Camera	Realtime Crime Detection and Tracking	High Resolution IP Camera with 20X Optical Zoom	B&H Photo Video	\$2,844.87
Nikon Lens		High quality fixed focus lens for 4K Camera	Edwin Watts Golf	\$399.99
MODULE 802.15.4 1MW W/U.FL CON			NewEgg.com	\$45.59
(2) Samsung 840 EVO SSD		High capacity solid state drive for high data rate video acquisition	Edmund Optics	\$679.98
Prosilica gige camera 63-756			Edmund Optics	\$107.99
Prosilica gige camera and power			Amazon.com	\$217.99
16GB Notebook Memory Modules	UAS Dataset Collection	Computer memory	Newegg.com	\$145.99
SSD 500G Internal Solid State	UAS Dataset Collection	High speed storage to enable realtime vision applications	Newegg.com	\$300.98
Gigabyte Brix Compact PC	UAS Dataset Collection	High performance small form factor computing platform for realtime visions applications	Newegg.com	\$509.99
<b>UAS Accessories</b>				
Cinemizer OLED Goggles	UAS Dataset Collection	Virtual reality headset for live video feedback from aerial platforms	GetFPV.com	\$814.32
Multiplexer Radio Controller	UAS Dataset Collection	Switching hardware that enables transitioning from autopilot control to remote user control	Acroname	\$39.95
Futaba Computer Radio and Receiver	UAS Dataset Collection	14 Channel radio control systems for UAS operation	Tower Hobbies	\$779.94

Long Range RC Transmitter	UAS Dataset Collection	UHF Transmitter to allow beyond line of sight UAS control	Dagon Link Labs	\$268.00
---------------------------	------------------------	---	-----------------	----------

## References

- [1] Haroon Idrees, Imran Saleemi, Cody Seibert, Mubarak Shah, Multi-Source Multi-Scale Counting in Extremely Dense Crowd Images, Computer Vision and Pattern Recognition (CVPR), Portland, Oregon, 2013.
- [2] P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan, Object Detection with Discriminatively Trained Part Based Models, IEEE PAMI, 2010.